Virtual Institute of Microbial Stress and Survival
DOE Genomes To Life Project
Progress Report: August, 2003

I. Overview

The objective of this monthly progress report is to provide an update of the technical and administrative actions from the previous month as well as forecast upcoming progress for the VIMSS Genomes to Life Project. I want to remind everyone how important to make sure everyone is communicating. The discussion boards (http://genomics.lbl.gov/~aparkin/discus) provide a forum for people to ask questions about direction of the project, priorities, and technical issues that can be read and answered by the entire group. I know email is often the most efficient means but it does privatize some of the important communications. Also, posting project data and information to BioFiles (https://tayma.lbl.gov/perl/biofiles) is EXTREMELY important. We are in the process of adding user help files to BioFiles – if you have user questions, please contact Keith Keller (tel: 510.495.2766 or email: kkeller@lbl.gov). This is the best metric I can give to the DOE leadership that we are making progress aside from the VIMSS website. Please make us and yourselves visible by donating data and information to the website.

II. Applied Environmental Microbiology Core

LBNL

SR-FTIR. We continue to modify the existing SR-FTIR spectromicroscopy apparatus to study *Desulfovibrio vulgaris* under anaerobic conditions. We continued to develop protocols for producing *Desulfovibrio vulgaris* biofilm on reference surface and on different types of mineral surfaces. We continued to establish the IR spectral baseline of *D. vulgaris* biofilm under anaerobic conditions. Validated IR results by live-dead stain fluorescence microscopy. We succeeded in (1) developing one of a series of protocols to produce *D. vulgaris* biofilm on a reference surface, (2) obtaining the IR spectral baseline of *D. vulgaris* biofilm under anaerobic conditions, and (3) confirming the IR results by means of live-dead stain microscopy. Used the anaerobic facility in Holman's lab and the SR-FTIR spectromicroscopy established at ALS and repeated the earlier experiment to achieve the goals.

The main focus of this month was producing growth curves with the new defined medium to prepare for biomass production in September. The first part of the month was spent obtaining the new chemicals after the new protocol for the LS3D medium was proposed. Several growth curves were run, indicating that the lag time for the new medium was 35 to 40 hours long. After consulting with Judy Wall's group it was discovered that the lag time could be shortened to 20 hours if vitamins were prepared in

the laboratory using Thauer's recipe rather than using premixed ATCC Wolfe's vitamins. The only difference between the vitamins is that in the Thauer's there is 10x greater vitamin B12 and one additional component: choline. The new medium formulation using Thauer's vitamins was named LS4D. The 20-hour lag time is still definitely a concern because this indicated that the cells have to overcome some stress before entering in log phase. Medium modifications are being tested to try to understand what is causing the stress. One growth curve was run with 10x the amount of vitamins outlined in the protocol. This did not shorten the lag time, but it did result in a higher OD during stationary phase. OD measurements are shown below both on semi-log and standard plots for some of the growth curve attempts. From our measurements, lag phase cell densities are on the order of 10^7 cells/ml, and stationary phase are about 10^9 cells/ml.

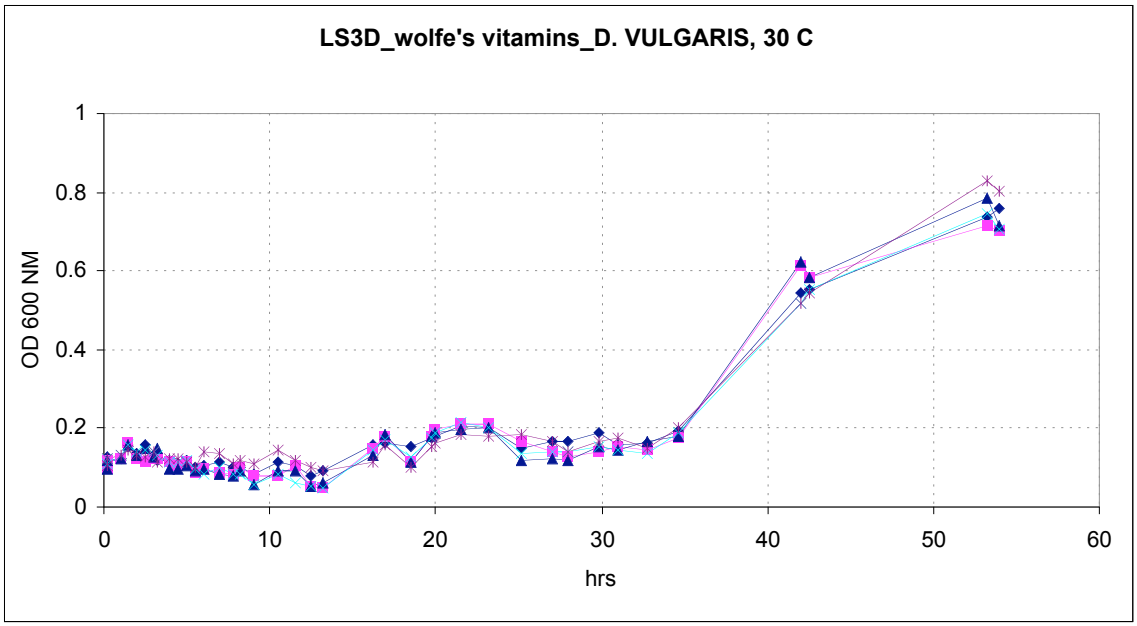
During growth curve measurements, modifications were made to the bottles and sampling system to assure the cultures in the bottles were pure and anaerobic. The resazurin in the medium proved invaluable to determine if there were any air leaks.

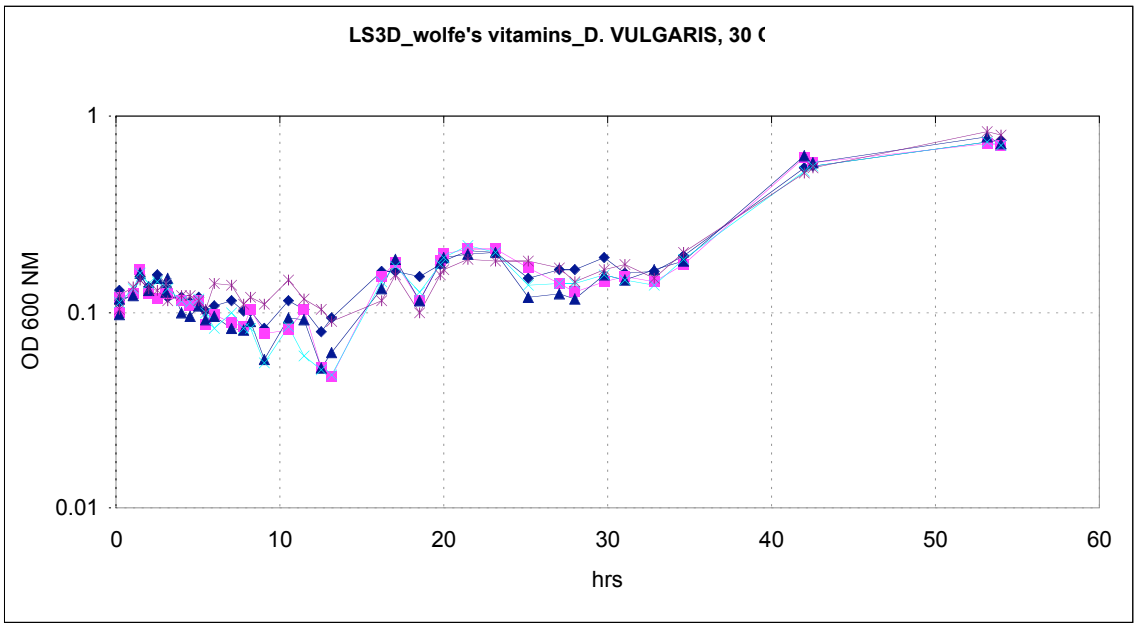
Some issues have arisen in the methods used to quantify biomass. OD 600 nm measurements are proving very problematic, because of changes that occur in the background medium over time due to the reductant, the Fe sulfides, and the resazurin. Two types of samples are used to blank the spectrophotometer: (1) excess medium stored anaerobically either in the anaerobic hood or in hungate tubes and (2) filtered culture. Depending on the growth conditions, these can give similar numbers for OD, or very different values. Additionally, when the air sparge starts, the bottles turn pink and the OD jumps up significantly due to the color change (e.g. from 0.3 to 0.7). Spectral scans have been run to try to find a better wavelength for quantification, but more work needs to be done.

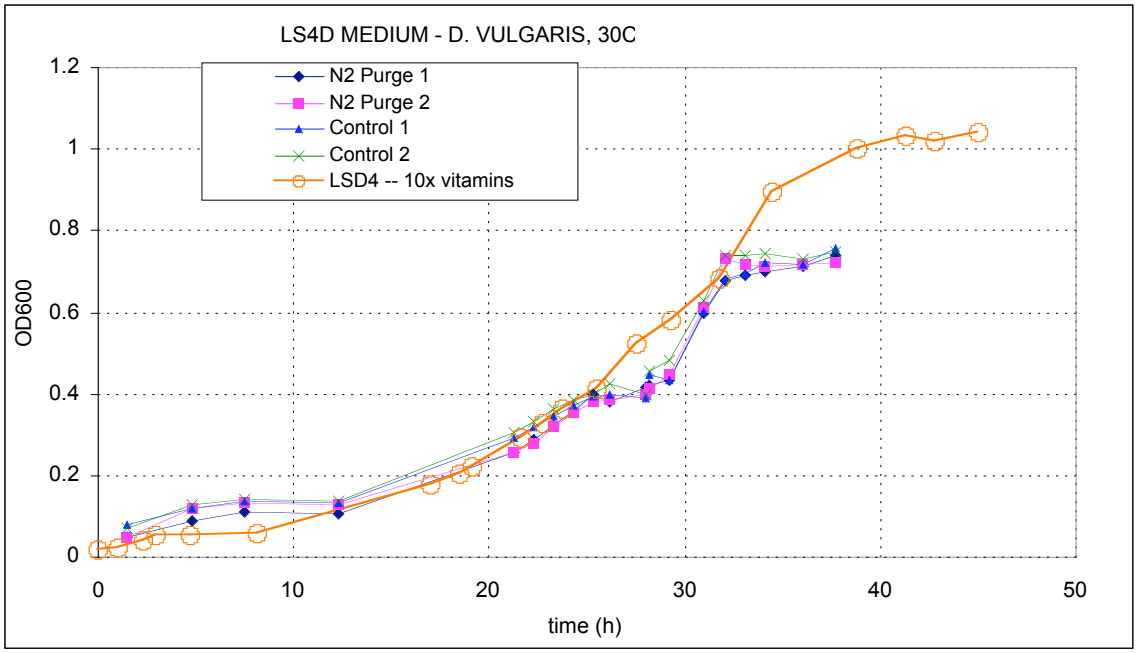
Cell counts have been used to verify purity and to quantify cell growth. Tween 80 and sodium pyrophosphate are added to the cells and they are vortex vigorously before mounting on the slides. However, significant clumping still occurs which is lowering the accuracy of the counts. New protocols are being developed to break apart the cells.

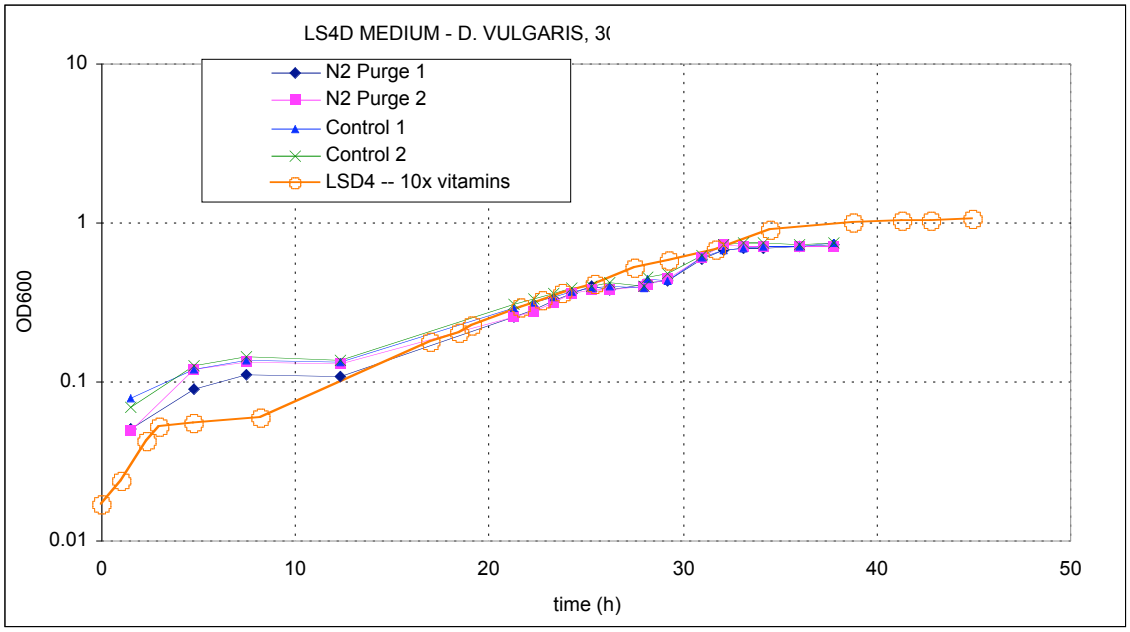
We also have found that the DC and the Bradford protein assay quantification ranges are high for the cell densities we are finding, especially during the lag phase. There are other commercially available assays and we will evaluate their efficacy in reducing the noise level in our protein data.

The protocol has been modified and cell production will begin in early September.


LS3D_wolfe's vitamins_D. VULGARIS, 30 C
OD 600 NM
1
0.8
0.6
0.4
0.2
0
0
10
20
30
40
50
60
hrs


LS3D_wolfe's vitamins_D. VULGARIS, 30 C
OD 600 NM
1
0.1
0.01
0
10
20
30
40
50
60
hrs





University of Washington

Optimization of Syntrophic Co-cultures. We have continued to optimize conditions for syntrophic growth of two alternative *Desulfovibrio* strains (Hildenborough and PT2) in co-culture with a hydrogenotrophic methanogen, Methanococcus *maripaludis*. A modified McC medium containing a reduced concentration of NaCl (2g/L) was used to investigate the influence of agitation and ratio of headspace to culture medium volume on growth rate. Two different head space/medium volume ratios (1/4 and 4/1) in

combination with different agitation regimes (designed to alter mass transfer between gas and liquid phases) were examined, monitoring optical density and methane production. Two variations of the modified McC media were also evaluated; one amended with 50 mM lactate, 80%/20% N_2/CO_2 head space, no sulfate and another amended with 50 mM lactate, 80%/20% N_2/CO_2 head space plus 2mM of acetate. The fastest syntrophic growth was observed using the small head space/medium volume ratio without agitation. In addition, a significant difference in growth rates of two co-culture was observed, the Hildenborough-methanogen co-culture had a doubling time of approximately 200 hours (Fig. 1A, B.) in contrast to less than 100 hours for the PT2-methanogen co-culture (Figs. 1C, D).

The co-culture of PT2 and the methanogen produced more methane than the Hildenborough co-culture under all growth conditions evaluated. Although the PT2 co-culture in tubes with a high head space/medium volume ratio (4/1) demonstrated a slightly higher growth rate and biomass yield, methane accumulation was much lower (about ten times) than in the tubes with a reduced head space/medium volume ratio (Fig. 1C, D; Fig. 2C, D). No methane accumulation was observed in co-cultures of Hildenborough with *M. maripaludis* in tubes with a 4/1 head space/medium volume ratio. Acetate (2mM) did not significantly influence either growth rate or methane production; only a slight increase in methane accumulation was observed for the co-culture of PT2 with acetate addition (Figs. 1 and 2).

Culture-based characterization and isolation of sulfate-reducing bacteria from the FRC. MPN enrichments of sediment sample FWB203-03d 04 from FRC area 2 on lactate, acetate, propionate pyruvate, ethanol or H2:CO2 showed sulfate reducing activity. T-RFLP analysis and cloning of PCR amplified 16S and *dsrAB* genes are now being used to characterize the bacterial composition of these enrichments.

Several bacterial colonies from selective agar amended with lactate or acetate with H2:CO2 as electron donors inoculated with FWB203-03d-04's MPN enrichments were picked and transferred into fresh B3 agar medium for pure culture isolation.

The fixed bed bioreactor systems were reconfigured to maintain anoxic conditions and inoculated with a pure culture *D. vulgaris* Hildenborough. Currently we are monitoring the process of biofilm formation in the reactor.

The *dsr*AB genes were cloned from three of eight sulfate-reducing bacteria isolated from the heavy metal impacted sediments of Lake Depue Previous analyses of their 16S rRNA genes demonstrated 99% identity to *D. vulgaris* Hildenborough. Consistent with previously determined relationship of the 16S rRNA genes, initial comparative analyses reveal greater than 99% sequence identity among *dsrAB* genes.

Immediate future work. Continue optimization of syntrophic growth of the two *Desulfovibrio* strains with the methanogen, examining

1. Stability of association in several consecutive passages

2. Monitoring of metabolite consumption and evolution and methane production.

3. Complete dsrAB gene sequence analyses of *Desulfovibrio* strains isolated from Lake Depue metal-contaminated sediments.

4. Complete dsrAB gene sequence analyses of *Desulfovibrio* strains isolated from Lake Depue metal-contaminated sediments.

5. Amplify and sequence 16S-23S intergenic spacer region of *Desulfovibrio* strains isolated from Lake Depue.

6. Continue isolation and characterization sulfate-reducing bacteria from FRC enrichments.

7. Amplify and sequence 16S rRNA and *dsrAB* genes from selected isolates.

The fixed bed reactor will be tested for its ability to support and maintain the growth of *Desulfovibrio* strains as biofilm monocultures.

Fig. 1. Growth of co-cultures DvH and PT2 with methanogen

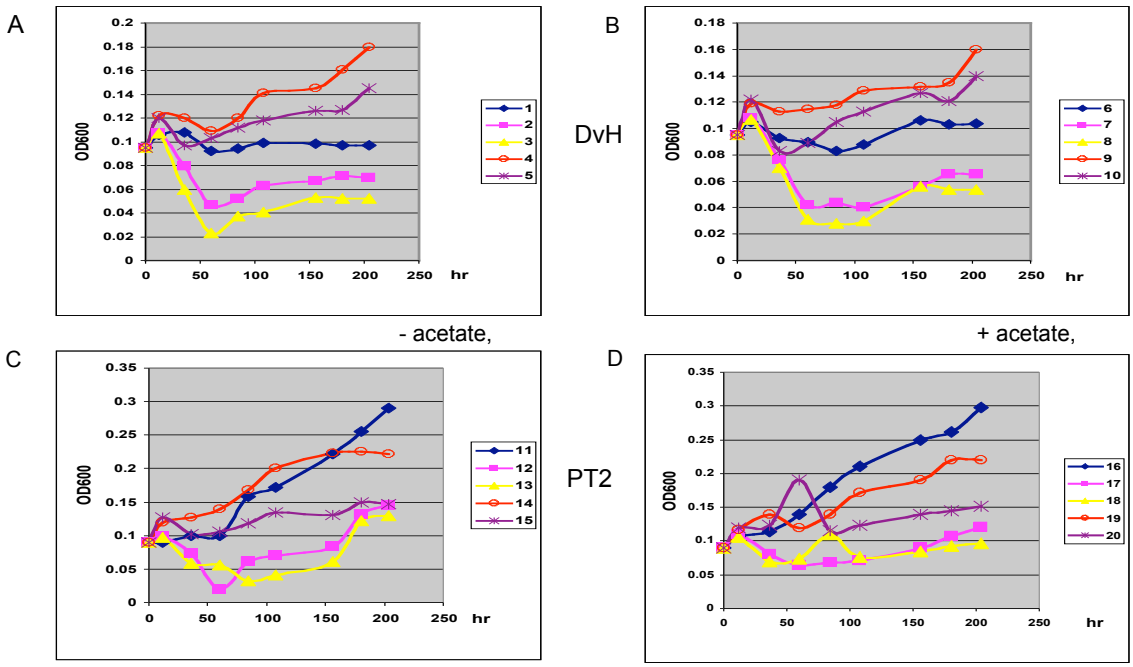
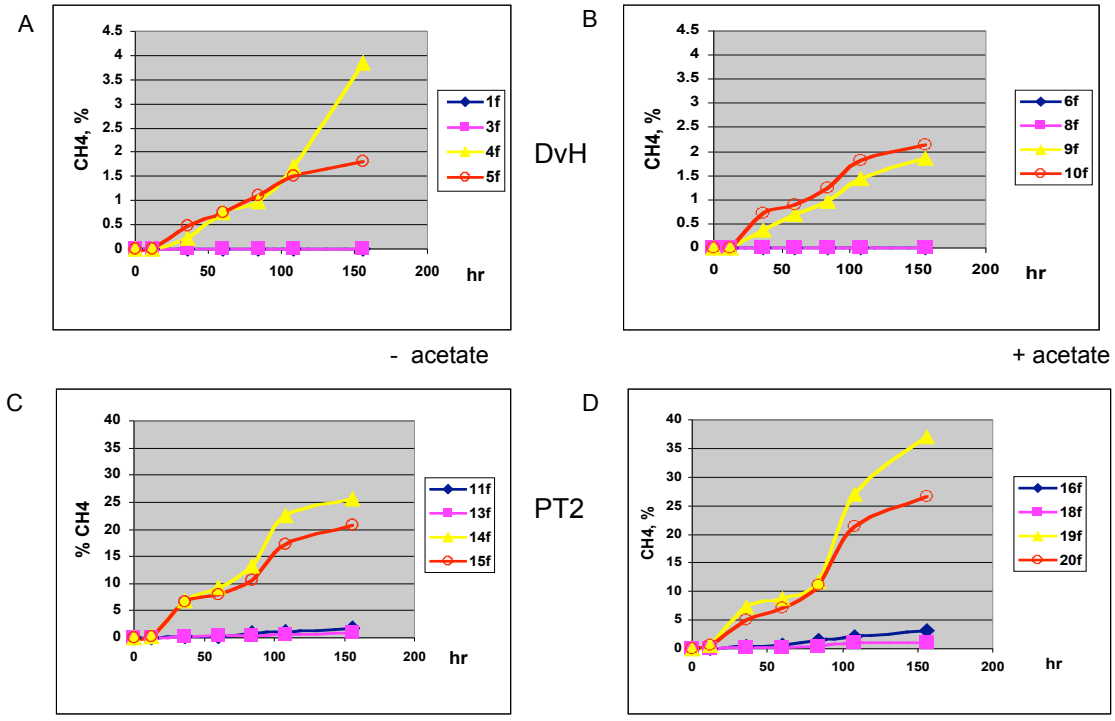


Fig. 2. Methane accumulation during growth of co-cultures DvH and PT2 with methanogen



Figs 1 and 2 legend: 1,6 – DvH+Mm, 4/1 head space/ medium ratio, no agitation; 11,16 – PT2+Mm, 4/1 head space/ medium ratio, no agitation; 2, 7 - DvH+Mm, 4/1 head space/ medium ratio, moderate agitation; 12, 17 - PT2+Mm, 4/1 head space/ medium ratio, moderate agitation; 3, 8 - DvH+Mm, 4/1 head space/ medium ratio, vigorous agitation; 13, 18 - PT2+Mm, 4/1 head space/ medium ratio, vigorous agitation;m4, 9 - DvH+Mm, 1/4 Head space/ medium ratio, no agitation; 14,19 - PT2+Mm, 1/4 head space/ medium ratio, no agitation; 5, 10 - DvH+Mm, 1/4 Head space/ medium ratio, vigorous agitation; 15, 20 - PT2+Mm, 1/4 head space/ medium ratio, vigorous agitation.

Oak Ridge National Laboratory

Nothing new this month.

Diversa

Progress:

- Genomic DNA from extremely dilute noodle extractions was concentrated,

amplified, and sent for diversity indexing and large insert library construction.

- Genomic DNA amplification reactions continue to be optimized for product size and genome coverage.

- Small insert libraries have been constructed for all current GTL soil samples, and are ready to be screened.

Issues:

- Screening of libraries is dependent on selection of a suitable target.

Actions:

- High GC random primers have been used to amplify *S. diversa* genomic DNA, which is GC rich (~70%). This DNA will be hybridized to an Affymetrix GeneChip to look at genome coverage.

- Work is ongoing to optimize the large insert FACS biopanning protocol in gel microdroplets. Currently, experiments are in progress to test the level of sensitivity of the assay.

- Plans are in progress for Diversa employees to visit LBNL and extract DNA from U-238 contaminated environmental samples.

III. Functional Genomics Core

Transcriptomics (ORNL)

Objectives

- To perform initial microarray expression profiling studies for the model bacterium *S. oneidensis* MR-1 to establish a baseline response to various environmental stresses.

- To create and use whole-genome microarrays for *Desulfovibrio* and *Geobacter* for analysis of stress-induced transcriptomes.

Progress since last report

- Microarray data on pH stress response in *Shewanella* has been analyzed in terms of distributions of log ratio expression difference for gene pairs within the same operon or selected at random. In addition, the GTL Computational Core Group has compared *S. oneidensis* and *E. coli* orthologous gene expression for pH stress. Writing of the manuscript for this project is in progress.

- GTL Computational Core has also completed operon analysis for the salt stress and heat shock data.

- The tolerance of *S. oneidensis* MR-1 to high levels of Sr^{2+} (180 mM) was examined using whole-genome DNA microarrays. A comparison of the transcriptomes generated from the strain growing in the presence or absence of Sr^{2+} revealed significant differences in the gene expression profiles. Microarray analyses revealed that a number of genes for siderophore biosynthesis and iron uptake were highly induced (6.5 to 118 fold) in the divalent metal-ion tolerant state. Several open reading frames predicted to encode hypothetical proteins were also highly induced (130-132 fold) in cells exposed to high levels of Sr^{2+}, whereas genes that encoded ATP synthase subunits were repressed (~3 fold) in the presence of the metal-ion. In addition, an insertional mutant for one of these siderophore biosynthesis genes (SO3032) shows reduced tolerance to Sr compared to the wild type.

Future work

- Complete pH stress manuscript.
- Continue microarray expression profiling of MR-1 cells exposed to toxic levels of certain metals (e.g., strontium), examining both shock and acclimation. Investigate tolerance of *Shewanella* to other metals (e.g., chromium, cobalt, arsenate)
- Target more genes for deletion mutagenesis based on expression data.
- Start global gene expression profiling for *Desulfovibrio*.
- Complete *Geobacter* whole-genome microarrays.

Proteomics (Diversa)

Objectives

- The activities for this month are mainly in the areas of fine tune of chemical and enzymatic reactions and data analysis.

Progress since last report

- Tested the approaches of improving the specificity and efficiency of the reduction/alkylation and digestion procedures. (See Table I & II).
- These samples are being analyzed. Results are to be determined.
- Progress was made in the conversion of LC-MS/MS data acquired by Q-STAR mass spectrometer. The remaining issues are the merge of duplicates MS/MS spectra prior to peak centroiding and the suitable parameters for peak centroiding.
- Similarly, efforts went into upgrading the data conversion program for LCQ data that requires new version of the library. We're in the process of getting help from Thermo-Finnigan on some software compatibility problems.
- Validation of the component detection continues using the data from 1D LC-MS/MS experiment. The accuracy and robustness were further improved.

Table 1. The composition of the protein mixtures and the concentrations of the constituents. The total concentration is 1 mg/ml.

BSA (2mg/ml)	40 ml	80 mg
a-casein (2mg/ml)	80 ml	160 mg
b-lactoglobulin (2mg/ml)	160 ml	320 mg
Cytochrome c (2mg/ml)	320 ml	640 mg
[Arg8]-vasotocin (1mg/ml)	40 ml	40 mg
Neuromedin B (1mg/ml)	40 ml	40 mg
H_2O	600 ml	
Total	1280 ml	1280 mg

Table 2. The matrix of the conditions for the optimization of reduction / alkylation procedure.

Samples	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	
	UDI	*RDI*	*UD6*	*RD6*	*UTI*	*RTI*	*UT6*	*RT6*	
Conc (mg/ml)	1	1	1	1	1	1	1	1	
Protein (mg)	(50)	(50)	(50)	(50)	(50)	(50)	(50)	(50)	
Vol (ml)	50	50	50	50	50	50	50	50	
2 M urea (ml)	50	-	50	-	50	-	50	-	126 mg / 1.05 ml
1% Rg	-	50	-	50	-	50	-	50	2 mg → 0.2 ml
100 mM DTT (ml)	1	1	1	1	-	-	-	-	→1 mM, 37C, 30min
100 mM TECP (	-	-	-	-	1	1	1	1	→1 mM, 37C, 30min
100 mM IAM (m	2	2	-	-	2	2	-	-	→2 mM, 37C, 30min
100 mM D6 (ml)	-	-	2	2	-	-	2	2	→2 mM, 37C, 30min
0.5 mg/ml Trypsin (ml)	1	1	1	1	1	1	1	1	→4:00 pm, 37C, o/n
	1	1	1	1	1	1	1	1	→10:00 am, 37C, 6h
Formic acid	-	0.6	-	0.6	-	0.6	-	0.6	→37C, 45 min
Final vol (ml)	106	106	106	106	106	106	106	106	
Conc (mg/ml)	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	
100 mM DTT (MW 154)	1.54 mg/100 ml								
100 mM IAM	1.85								

(MW 185)	mg/100 m								

DTT= dithiothreitol
TCEP= tris(2-carboxyethyl)phosphine
IAM=iodoacetamide
D6=proprietary alkylating reagent (similar function as IAM)

Protein complexes (Sandia)

On-going work

- Isolation of Protein Complexes:
  Further obtained whole cell protein extract for stress conditions of UV for *E. coli* and heat shock for *D. vulgaris*. Currently preparing samples to be run through the DIGE process for replicating previous results. Prepared antibody coupled gels (RecA and Hsp70) for Co-IP runs of the aforementioned whole cell protein extracts.

- Mass Spectrometry for Protein Complex Identification:
  NanoLC setup completed. Began testing columns, gradients, and ion pairing reagents with protein digest standards to optimize separation conditions. Thus far achieved 20 discernable peaks in the UV chromatogram of a BSA tryptic digest (~50 peptide components).

Future Work (short term)

- Replicating DIGE runs for heat stress in *D. vulgaris*
- Performing DIGE runs for oxygen stress in *D. vulgaris*
- Setting up MALDI instrumentation (Voyager DE PRO - Applied Biosystems) for complex identification.

Metabolomics and Proteomics (UCB, LBNL)

Objectives

- Determine the detection limits of nucleotides analysis by ESI-MS-MS.
- Determine the range of linear response for nucleotides analysis by ESI-MS-MS.
- Establish a list of expected metabolites in *D.vulgaris* based on KEGG metabolic pathways.
- Determine antibiotic concentrations effective for *D. vulgaris* growth on plates.
- Complete a successful conjugation into *D. vulgaris*.
- Construct gene knockout vector for large scale inactivation of *D. vulgaris* genes

- Work out methods for cell lysis of *D. vulgaris* using French Press, and attempt to quantify differences between proteomes isolated using this method and sonication. This is in preparation for the actual completion of an oxygen stress experiment.

Progress since last report

- Nucleotide separation conditions were further optimized to obtain better resolution.
- Detection limit of ATP, ADP AMP are 70 pmol, 11 pmol and 12 pmol respectively.
- The dynamic range of ATP, ADP and AMP is up to 2.5 pmol.
- A CE-UV-MS system was set up and ready for method development.
- Based on the KEGG biochemical pathway maps predicted from the genome sequence of *D. vulgaris*, the Computational Genomics group (Katherine and Eric) generated a list of 618 metabolites from 403 reactions. This list was examined manually and reduced down to ~400 small organic molecules.
- Several steps have been successfully conducted to obtain required vector for the single cross-over knock-out strategy in *D.vulgaris*. The pUC19 vector was amplified sans its ampicillin resistance cassette. The chloramphenicol resistance cassette was amplified from pACYC184. Via hybridization using 15bp overhangs, the pUC19Cam vector was created. Next, pUC19Cam was amplified to generate an open vector with primers designed to leave out the *lacZ_ orf*. An *oriT* gene was amplified from pEX100T and hybridized with the open pUC19Cam. The resulting vector is CamR and produces white colonies. In both hybridizations high yields were observed by electroporation in DH10B cells, as reported earlier in Fknock out the final step, the *lacZIacI* gene has been amplified from pHRP308 and will be placed behind the P_{lac} promoter.
- Cell lysis, protein quantification and ICAT labeling methods were developed in anticipation of the expected biomass for the Oxygen stress experiment. It was observed that French Press Lysis using cells harvested from approximately 50ml culture in mid log phase (16 hours in LS media) and resuspended in 3ml Tris.HCL pH 8 provided a lysate with adequate quantity of protein (~ 5-10 ug/ul) concentration for the ICAT protocol.
- ICAT samples of a French Press prepared sample was used as a baseline for stress experiments.
- A conjugation experiment of *D. vulgaris* was completed on Postgate media, following Voordow's protocol, however it was unsuccessful.

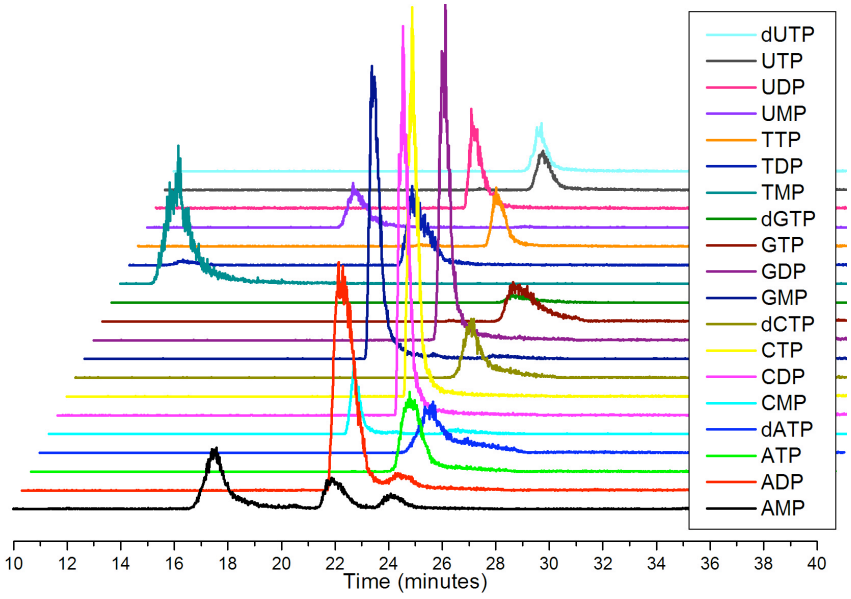


Figure 1. Separation of nucleotides. A hydrophilic interaction column was used. The chromatographic conditions are: 10 minutes elution by 20% buffer (NH_4OAc, 5 mM, pH=5.0) + 80% ACN, followed by a linear gradient to 60% buffer in 20 minutes. The signals were detected by a QTrap in a positive MRM mode. The ion source temperature is 400 ^oC.

Future work

- Continue the quantization of nucleotides
- CE method development for amino acids, pentose phosphate pathway, glycolysis pathway and TCA cycle intermediates.
- Determine antibiotic concentrations effective for *D. vulgaris* growth on plates. We will be using LS3 media for our work from now on.
- The metabolite list which includes MW and category will be further examined against the KEGG pathways to determine if some metabolites were omitted. The final list will be posted on the VIMSS web site.

IV. Computational Core

LBNL – Arkin

In August, we released the VIMSS Comparative Genomics Website to GTL collaborators. We purchased and installed a 4-node webserver to host the site in anticipation of high traffic following the public (non-GTL) release of the site, which will accompany the publication of an applications note describing the VIMSS tools in a peer-reviewed journal. The VIMSS tools include the Comparative Genomics browser, protein pages, the VIMSS Regulon Browser, sequence alignments to common protein domains and a browser for VIMSS operon predictions.

We worked closely with the Functional Genomics Core in August on three projects. To assist Vince Martin in metabolomic analysis, we set up tools to predict and to browse metabolic pathways in our three target organisms. We are working with collaborators at ORNL to analyze the Geobacter metallireducens sequence to determine the genes for which oligos are still needed for microarray analysis. In collaboration with the Zhou laboratory at ORNL, we have started comparative analysis of microarray data from different species to elucidate common regulatory schemes, as well as using our operon and regulon predictions to help assess experiment quality.

In August, we finalized the gtl mailing lists and reconfigured the webserver to minimize downtime during updates. In addition, some progress has been made on extending Biofiles functionality. In particular, we have begun work on a sample tracking module, and during the next month hope to begin work on storing uploaded data in a custom schema such that we can offer more advanced tools for querying and viewing data from both the Applied Environmental and Functional Genomics Cores. In addition, this month we developed a schema to represent data from BioCyc in a relational (MySQL) database, and have integrated these data into the VIMSS Comparative Genomics Database.

Over the next month, in addition to continuing development on these projects, we plan to install a new fileserver to accompany our move to new office space, hire and train a new programmer, finalize gene models for Geobacter metallireducens for ORNL, and complete a substantial update of new sequence to our Comparative Genomics Database.

LBNL – Olken

Visual Graph Query Language

Work continued on design of a visual graph query language user interface. We have concluded that we will need a multi-paradigm UI. Graph matching queries will use a query graph drawing paradigm (nodes, edges, paths). Other queries – Boolean graph operators, aggregation, clustering, etc. will employ an operator tree paradigm (e.g., similar to logic diagrams). Contextual queries will use a textual (logic) specification for the context. Selection from various taxonomies will use a series of cascading menus (like directory trees) to navigate through the taxonomies.

Navigational API

Work continues on the design of the navigational API.

Plans for September

1. Continue development of RDF/web based schema browser prototype.
2. Continue development of RDF/web based relational DB instance browser.
3. Complete biopathways chapter for DOE Computational Biology Primer.
4. Complete report on NLM Workshop on Data Management.

5. Complete Navigational API specification.
6. Code Navigational API.
7. Begin efforts to develop a common biopathway data interchange format with other members of Arkin Lab, the Ozsoyoglu's at Case Western Reserve Univ, Peter Mork (Univ. of Washington) and biopax.org.
8. Begin efforts to develop a common biopathway data interchange format with other members of Arkin Lab, the Ozsoyoglu's at Case Western Reserve Univ, Peter Mork (Univ. of Washington) and biopax.org.

V. Project Management

Project Schedule

The GTL project schedule is undergoing updates and will be posted to the VIMSS Discussion Board on a regular basis. Any updates/comments/revisions should be sent to Nancy Slater via email (naslater@lbl.gov).

FY04 Budget

The FY04 budget is under development. The DOE has not given any guidance regarding the status of the budget. A Continuing Resolution budget is anticipated until the federal budget is approved.